OSG Engage

Life on campus for individual and small team researchers

A Science Highlight: Steffen Bass

John McGee, Jason Reilly - RENCI
Mats Rynge - USC ISI

Where do researchers go for services?

- PI owned and operated cluster
- Campus Condominium Computing
- Departmental Cluster
- Campus Research Computing
- Campus Condor Pool

- State and Regional Initiatives (NYSGRID, NWICG, TIGRE)
- Communities of Practice (NanoHub, GridChem, NBCR, SBGrid etc)
- NIH Computational Centers
- TeraGrid: NSF, competitively awarded allocations
- Open Science Grid: DOE/NSF, opportunistic access
- DOE ASCR: INCITE awards
- Commercial Cloud service providers

Where do researchers go for services?

answer: wherever they can get them, with the least amount of pain

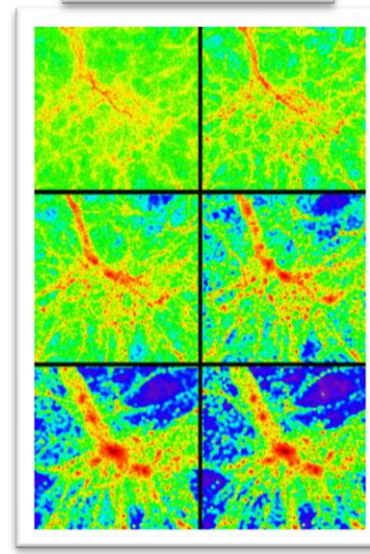
- PI owned and operated cluster
- Campus Condominium Computing
- Departmental Cluster
- Campus Research Computing
- Campus Condor Pool

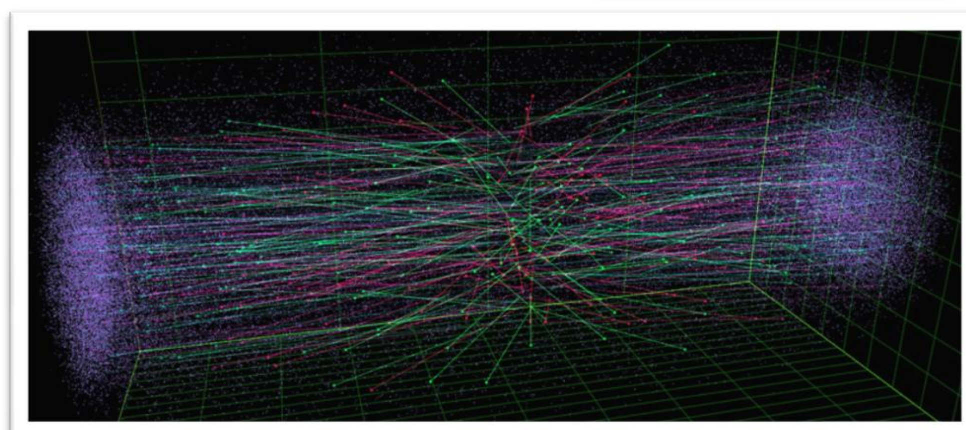
How many different:
service interfaces
software stacks
policy frameworks
identities per researcher
…
where is **The** National Cyberinfrastructure?

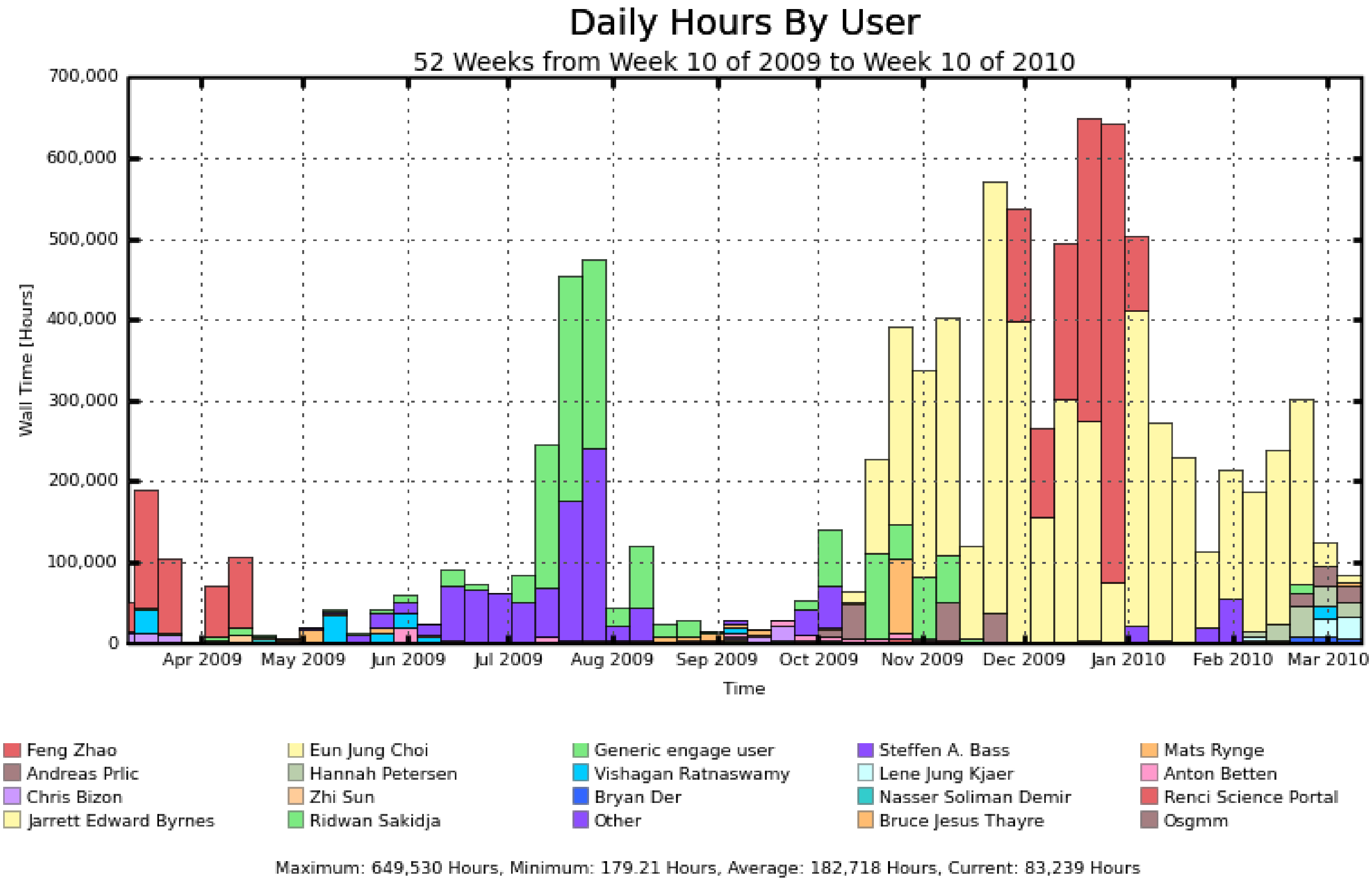
- State and Regional Initiatives (NYSGRID, NWICG, TIGRE)
- Communities of Practice (NanoHub, GridChem, NBCR, SBGrid etc)
- NIH Computational Centers
- TeraGrid: NSF, competitively awarded allocations
- Open Science Grid: DOE/NSF, opportunistic access
- DOE ASCR: INCITE awards
- Commercial Cloud service providers

OSG Engage Science Highlight: *Steffen Bass*

- Studies of quark-gluon plasmas: leading to better understanding of the beginnings of the universe
- Computational modeling: many complex models with many parameters
- New methods for constraining parameters and validating model assumptions
- Capabilities developed by this work **will revolutionize how simulations and data are analyzed**

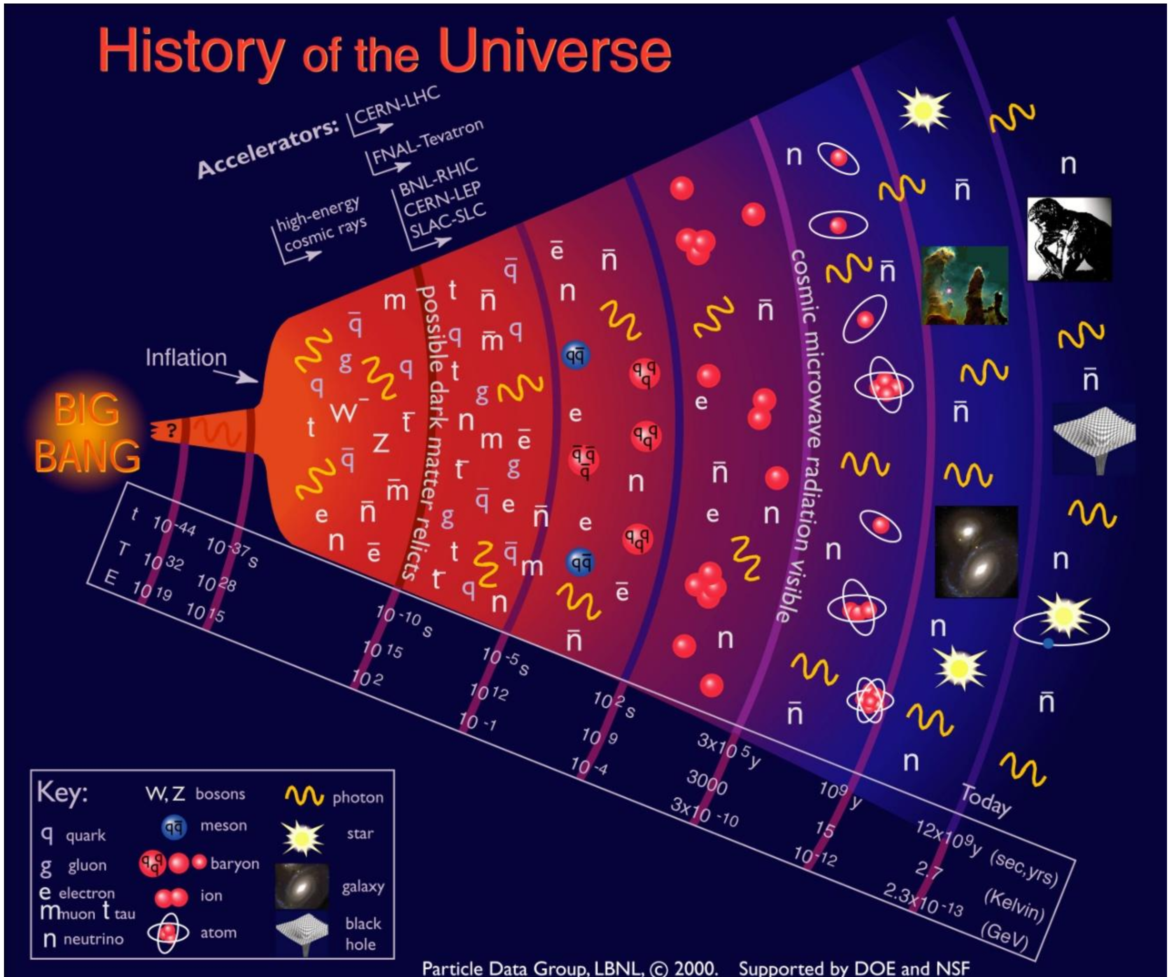






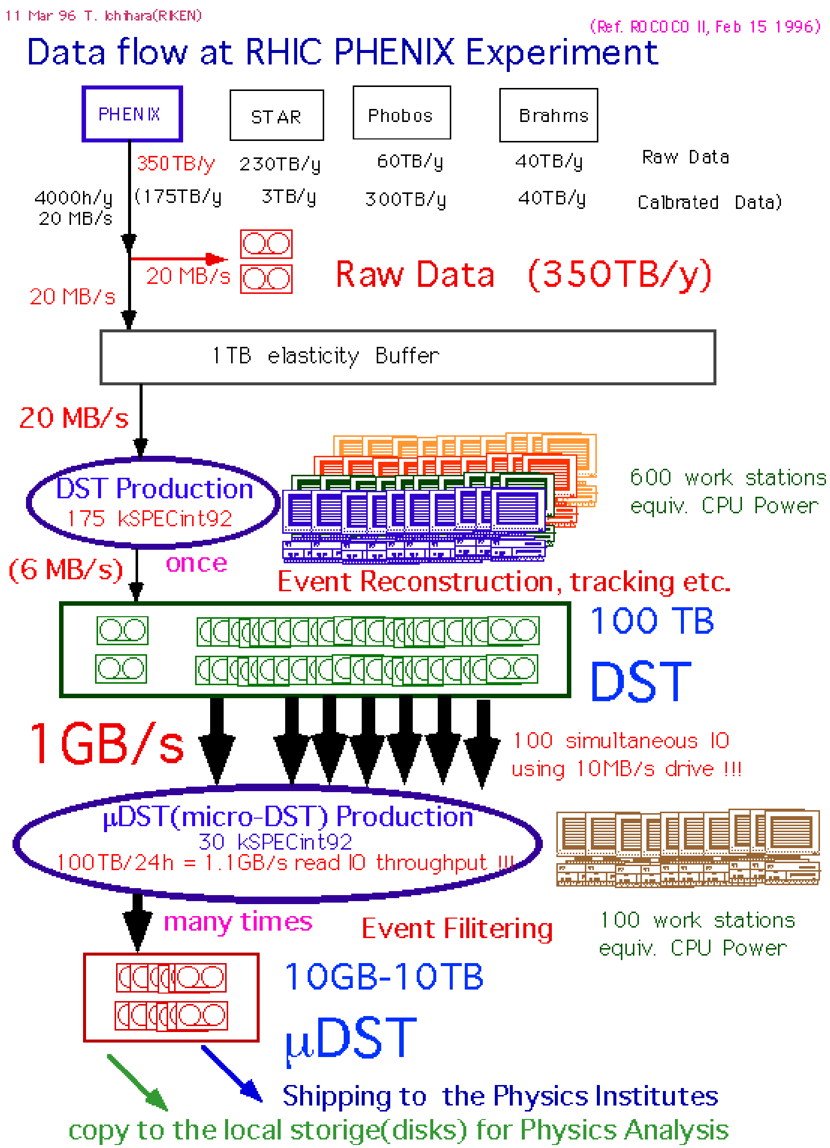
Daily Hours By User
52 Weeks from Week 10 of 2009 to Week 10 of 2010
Wall Time [Hours]
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0
Apr 2009
May 2009
Jun 2009
Jul 2009
Aug 2009
Sep 2009
Oct 2009
Nov 2009
Dec 2009
Jan 2010
Feb 2010
Mar 2010
Time
Feng Zhao
Eun Jung Choi
Generic engage user
Steffen A. Bass
Mats Rynge
Andreas Prlic
Hannah Petersen
Vishagan Ratnaswamy
Lene Jung Kjaer
Anton Betten
Chris Bizon
Zhi Sun
Bryan Der
Nasser Soliman Demir
Renci Science Portal
Jarrett Edward Byrnes
Ridwan Sakidja
Other
Bruce Jesus Thayre
Osgmm
Maximum: 649,530 Hours, Minimum: 179.21 Hours, Average: 182,718 Hours, Current: 83,239 Hours

The Quark-Gluon-Plasma: Exploring the Early Universe

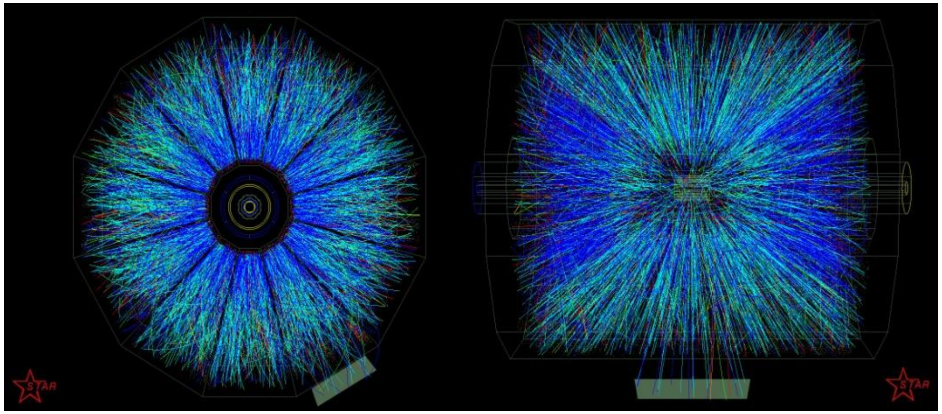


- the basic constituents of matter are quarks and gluons
- a few microseconds after the Big Bang the entire Universe was composed of a plasma of quarks and gluons (QGP)
- compressing & heating nuclear matter to a point where the nucleons dissolve into quarks & gluons allows to investigate the history of the Universe
- the only means of recreating temperatures and densities of the early Universe is by colliding beams of ultra-relativistic heavy-ions

RHIC Experiments & Data

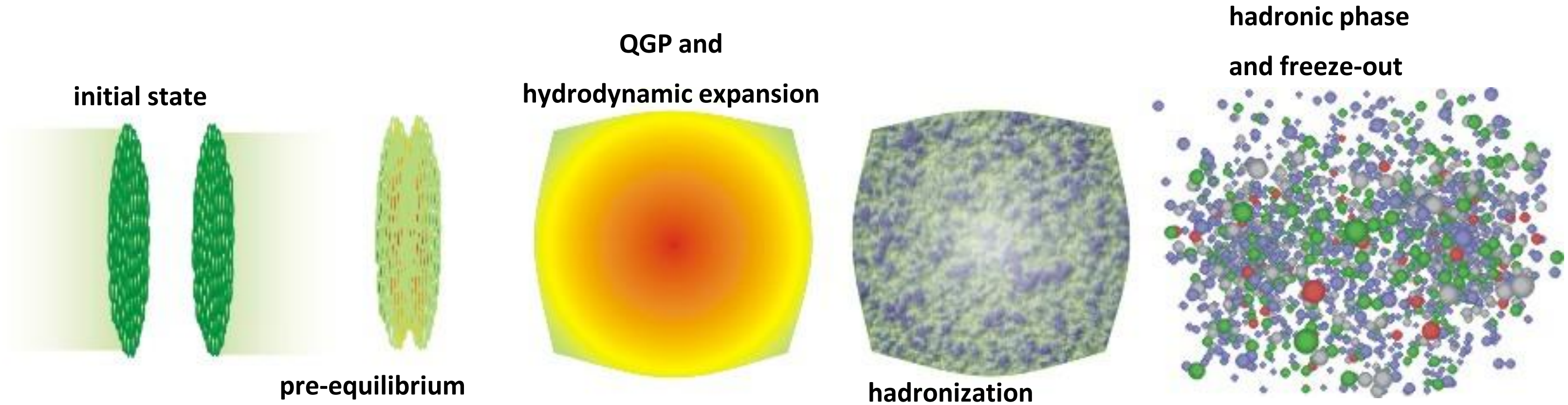


- typical collision recorded by the STAR detector: Au+Au @ 200 GeV/NN-pair
  - ➢ 1000s of tracks have to be reconstructed to determine species and momenta of produced hadrons and characterize collision



- several PetaByte of data have been collected since June 2000
- how to extract Physics conclusions from the collected data?

Knowledge Extraction: The Need for Modeling



Challenges:

•time-scale of the collision process: 10^{-24} seconds! [too short to resolve]

•characteristic length scale: 10^{-15} meters! [too small to resolve]

•confinement: quarks & gluons form bound states @ hadronization, experiments don’t observe them directly

Experiments:
- observe only the final state
- rely on QGP signatures predicted by Theory

Transport-Models:
- full description of collision dynamics
- connects intermediate state to measurements

Transport Models for RHIC

microscopic transport models based on the Boltzmann Equation:

- transport of a system of microscopic particles
- all interactions are based on binary scattering

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{E} \times \frac{\partial}{\partial \vec{r}}\right] f_1(\vec{p},\vec{r},t) = \sum_{processes} C(\vec{p},\vec{r},t)$$

diffusive transport models based on the Langevin Equation:

- transport of a system of microscopic particles in a thermal medium
- interactions contain a drag term related to the properties of the medium and a noise term representing random collisions

$$\vec{p}(t+\Delta t) = \vec{p}(t) - \frac{\kappa}{2T}\vec{v}\cdot\Delta t + \vec{\xi}(t)\Delta t$$

(viscous) relativistic fluid dynamics:

- transport of macroscopic degrees of freedom
- based on conservation laws:

$$\partial_\mu T^{\mu\nu} = 0$$

$$\begin{aligned} T_{ik} &= \varepsilon u_i u_k + P\left(\delta_{ik} + u_i u_k\right) \\ &- \eta\left(\nabla_i u_k + \nabla_k u_i - \frac{2}{3}\delta_{ik}\nabla\cdot u\right) \\ &+ \varsigma\,\delta_{ik}\nabla\cdot u \end{aligned}$$

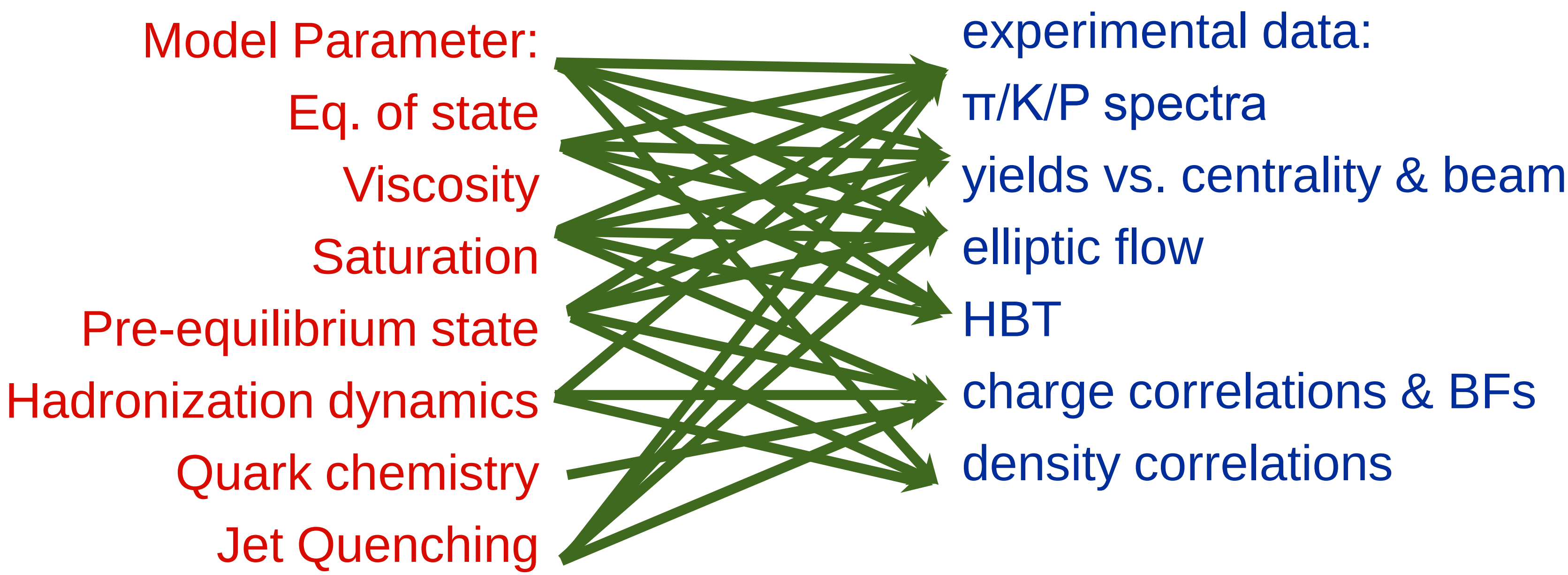
(plus an additional 9 eqns. for dissipative flows)

hybrid transport models:

- combine microscopic & macroscopic degrees of freedom
- current state of the art for RHIC modeling

Each transport model relies on roughly a dozen physics parameters to describe the time-evolution of the collision and its final state. These physics parameters act as a representation of the information we wish to extract from RHIC.

Making Connections: Pushing the Boundaries of Expertise



- large number of interconnected parameters w/ non-factorizable data dependencies
- data have correlated uncertainties
- develop novel optimization techniques: Bayesian Statistics and MCMC methods
- transport models require too much CPU: need new techniques based on emulators
- general problem, not restricted to RHIC Physics

→seek help/collaboration from Statistical Sciences

MaDAI Collaboration:
Models and Data Analysis Initiative

a multi-institutional and multi-disciplinary collaboration to develop next generation tools for complex model-to-data knowledge extraction

Michigan State University

RHIC Physics: Scott Pratt
Supernova: Wolfgang Bauer
Astrophysics: Brian O'Shea and Mark Voit
Atmospheric Modeling: Sharon Zhong
Statistics: Dan Dougherty

Duke University

RHIC Physics: Steffen A. Bass and Berndt Müller
Statistics: Robert Wolpert

UNC & RENCI

Visualization: Xunlei Wu and Russell M. Taylor

Funded by NSF CDI program (Cyber-Enabled Discovery Initiative)

- US$ 1,800,000 over 4 years

CDI: Extracting Science from Data & Models

- develop a comprehensive transport model (or set of consistent interlocking transport approaches), capable of describing the full time-evolution of a heavy-ion collision at RHIC, starting from the coherent glue-field dominated initial state up to the hadronic final state

- identify the relevant physics parameters (EoS, QCD transport coefficients, matrix elements etc.) which are sensitive to the observables measured at RHIC

- conduct a systematic study in that multi-dimensional parameter-space and via comparison to data to determine the properties of the QCD medium created at RHIC

Exploratory effort: understand how iRODS performs in managing Data between local campus storage system and NERSC archival allocation